

Genetic transition in the house mouse, *Mus musculus* of Eastern Iranian Plateau

Jamshid DARVISH¹, Annie ORTH² and François BONHOMME^{2*}

¹ Rodents Group, Ferdowsi University, Mashad, I.R. Iran

² Génome, Populations, Interactions, Adaptation, UM2 - IFREMER - CNRS UMR 5171, Université de Montpellier II, Place E. Bataillon cc 63, F-34095 Montpellier Cedex 5, France; e-mail: bonhomme@univ-montp2.fr

Received 16 February 2006; Accepted 13 November 2006

Abstract. Mice belonging to the *Mus musculus* species complex from the north-eastern Iranian Plateau (Khorasan province) have been genetically characterised for allozymic variation, mitochondrial DNA and Y chromosome type and compared with samples from other geographic regions. The present study shows the existence of a transition zone between pure *M. m. musculus* in the North and animals related to *M. m. castaneus* in the South. The origin of this transition (primary or secondary contact) and the various biogeographic scenarios about its origin are discussed in the light of these new data sets. The possible role of the Harirud valley in the geographic connection between Central Asia and the Middle East is discussed.

Key words: *M. m. musculus*, *M. m. castaneus*, biogeography, genetic transition, allozymes, mtDNA, Y chromosome, p53 pseudogene

Introduction

Among mammals, murids are a speciose family which tend to be complex from the biosystematical point of view. The house mouse *Mus musculus* is no exception and its division into a series of subspecies is now well documented (Boursot et al. 1993, Prager et al. 1996, Sage et al. 1993). Bonhomme et al. (1994) proposed that the Indian subcontinent was the cradle of this species and that a subsequent radiation led to the establishment of the well defined peripheral subspecies *M. m. musculus*, *M. m. domesticus* and *M. m. castaneus*. However the precise location of the cradle and the routes out of it is subject to debate. Prager et al. (1998) have for instance proposed that mice have colonised their present range in a more linear fashion, successively occupying the Middle-East and the Indian subcontinent from an ancestral population located in the Near-East. Milshnikov et al. (2004) proposed that the primary differentiation of *M. m. musculus* occurred further west in Transcaucasia. Hence, the systematic status of the intermediate populations from Eastern Iran, Afghanistan, Pakistan and Northern India (that we will hereafter refer to as “central”) remains unresolved and still subject to controversy. The poorly documented *M. m. bactrianus* from Afghanistan which is frequently cited in the literature provides a good example of this problem. Populations from regions adjacent to the putative centre of origin proposed by Bonhomme et al. (1994) are, therefore, interesting to study. In this context, Iran occupies an interesting position situated between the *musculus* populations in central Asia, the *domesticus* populations to the west and *castaneus* further east. Its border with Afghanistan occupies one of the possible passages towards Central Asia to the north, and in the south there is a passage along the coastal plains of the Persian Gulf that joins the Fertile Crescent to the Indian subcontinent.

*Corresponding author

Previous reports concerning the genetic characterisation of *Mus musculus* in this region are scarce. A mouse population trapped near Tehran in N Iran was shown to have nuclear genes that are intermediate between *M. m. domesticus* and those found in the Indo-Pakistani region (D i n et al. 1996), but its mitochondrial DNA lines all belonged to the Oriental clade defined by (B o u r s o t et al. 1996), which include the matriline found in, the peripheral subspecies *M. m. castaneus*. Other studies reported the presence of *domesticus*-type mitochondrial lines in samples from both the SE and SW Iran (G ü n d ü z et al. 2000). However in NE Iran (Khorasan province), B o i s s i n o t & B o u r s o t (1997) showed that the *musculus*-type mitochondrial lines come into contact with the oriental ones. In the north of the country, in the coastal plain located between Caspian Sea and the Elbourz mountain one *musculus* (Chalus) and two *castaneus*-type (Gorgan and Shahrahbad) matriline have been detected (P r a g e r et al. 1998). However, B o i s s i n o t & B o u r s o t (1997) only found *musculus*-type Y chromosomes in NE of Iran. The absence of pseudogene Ψ p53 reported by P r a g e r et al. (1998) also suggests the predominant presence of *musculus*-type genotypes in the north of the country. These results all point toward a complex biogeographic scenario in this region.

In the present report, we analysed allozymic, mtDNA, Y chromosome and Ψ p53 polymorphisms in three additional samples from the Khorasan province (Mashhad, Kakhk

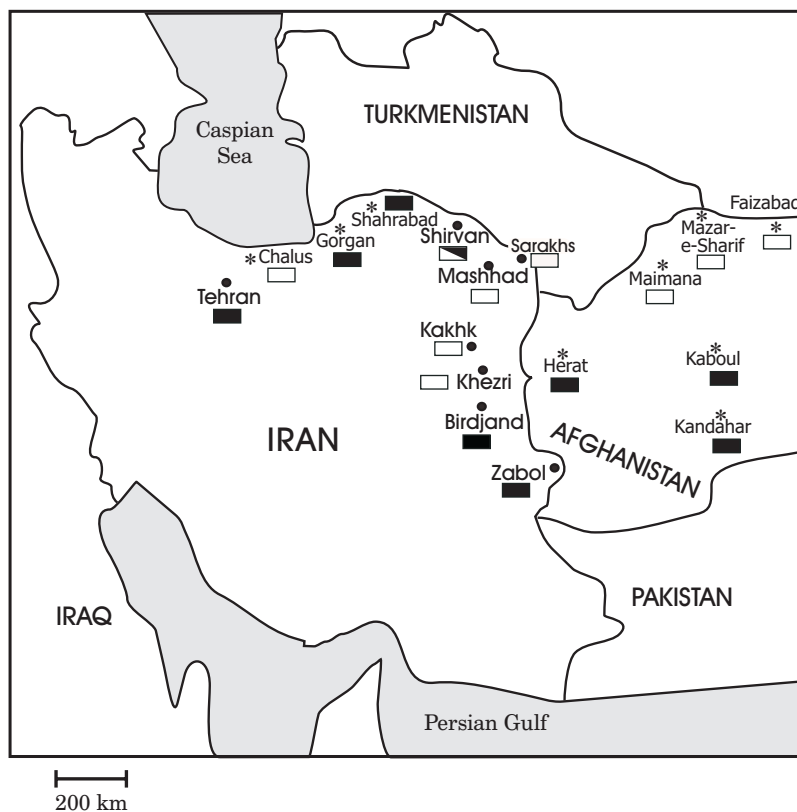


Fig. 1. Geographic locations of the sampled populations (dots). Asterisks indicate locations referred to in P r a g e r et al. (1998). Mitochondrial haplotypes in each locality are indicated as: ■ oriental mtDNA, □ *musculus* mtDNA, ▨ *musculus* & oriental mtDNA.

and Birdjand) and compared the results with the genetic information available in the literature. This allowed us to identify a transition from gene frequencies typical of *M. m. musculus* to those found in the populations from the Indian subcontinent. This supports the conclusions of a preliminary report on the morphological characterisation of samples from these two towns indicating that they could be assigned to two different entities (D a r v i s h 1995).

Materials and Methods

The localities sampled are shown in Fig.1. The origins of the mice analysed in this study are given in Table 1. Relevant reference samples from D i n et al. (1996) and B o u r s o t et al. (1996) were included in the analysis. The 17 polymorphic allozyme loci (see Table 2) were typed as described in D i n et al. (1996) and the Y chromosome gene *Zfy-2* as described by B o i s s i n o t & B o u r s o t (1997). The mtDNA of the individuals not included in the earlier studies were typed by PCR-RFLP for the presence/absence of the BamHI cleavage site at positions 3564, 9172 and 14232 of the mtDNA molecule (seq AY172335 (B a y o n a - B a f a l u y et al. 2003)), which yielded three different profiles: A P A for *castaneus*, A A A for *musculus* and P A A for *domesticus*. Primers used for these three loci were 3564: 5'AGCCTGACCCATAGCCATAA3' & 5'AAGACCGTTTGTGGTTGTTG3'; 9172: 5'CCACATAAATCAAGCCCTACT3' & 5'AGAAGGTAGATGGCATATTGGT3; 14232: 5'TATATACACGCAAACGGAGCC3' & 5'CTAGTTTATTGGGGATTGAGCC3'.

The presence/absence of the pseudogene Ψ p53 on mouse chr. 17 was revealed according to the experimental specification of P r a g e r et al. (1997).

Nei's genetic distance corrected for small sample sizes was computed from gene frequencies at allozyme loci (N e i 1978). The distance matrix was then transformed in distance trees using the Neighbour-Joining algorithm as implemented in Phylip 3.6 phylogenetic package (F e l s e n s t e i n 1993), and plotted using the TreeView program (P a g e 1996).

Results

Gene frequencies of 17 polymorphic allozymic loci are shown in Table 2. Fig. 2 represents the tree projection of the genetic distance matrix. As already reported in D i n et al. (1996) the Tehran sample falls in an intermediate position between the Indo-Pakistani populations and *domesticus*.

The three samples from NE Iran (Birdjand, Kakhk, Mashhad) we analysed for these allozyme loci are not only very different from the Tehran sample but are themselves genetically differentiated (Fig. 2). The sample from Birdjand is close to the Indo-Pakistani populations which group in the centre of the tree, while the sample of Mashhad in the north is almost pure *M. m. musculus*. The Kakhk is both geographically and genetically intermediate.

Table 1 reports the results obtained for the three DNA markers. It was possible to include four more localities (Sarakhsh, Shirvan, Kezhri and Zabol) in this part of the study. The Ψ p53 pseudogene was present in the two southernmost samples (Birdjand and Zabol), while it is absent in the northern ones (from Shirvan to Kezhri, Fig.1), which is characteristic of the *M. m. musculus* subspecies (O h t s u k a et al. 1996, T a n o o k a et al. 2001). This adds further support to the existence of a genetic transition between Mashhad and Birdjand

Table 1. Geographic coordinates of samples, references of data and results (DNA mt, ChY, Ψ_{p53}). Sample size (n). Sample reference (ref) : # This study; [a] (B o n h o m m e et al. 1984), [b] (B r i t t o n - D a v i d i a n et al. 1989), [c] (D i n et al. 1996), [d] (O r t h et al. 1996), [e] (B o u r s o t et al. 1996), [f] (B o i s s i n o t & B o u r s o t 1997); [g] (O r t h et al. 2002). The Ψ_{p53} results are tabulated as (P) for presence and (A) for absence of a *p53* pseudogene.

Country	Locality	Coordinates	Allozymes		mtDNA		ChY (Zfy-2)		Ψ_{p53}			
			n	ref	n	ref	n	ref	n	ref	type	type
Iran	Sarakhs	36.32N 61.07E	2	#	2	#	1	#	1	#	musc	N
Iran	Shirvan	37.25N 57.55E	4	#	4	#	1	#	4	#	musc	N
Iran	Mashhad	36.17N 59.33E	4	#	4	[f]#	3	[f]	4	#	<i>musc</i>	N
Iran	Kakhk	34.09N 58.38E	10	#	10	[f]#	2	[f]	2	#	musc	N
Iran	Khezri	33.52N 58.43E	3	#	3	#	3	#	2	#	musc	N
Iran	Birdjand	32.54N 59.11E	4	#	4	[f]#	2	[f]	2	#	musc	P
Iran	Zabol	31.15N 61.36E	4	#	4	#	1	#	4	#	musc	P
Iran	Tehran	35.41N 51.25E	8	[c]	8	[e] [f]#	1	[f]	2	#	<i>musc</i>	P
India S.	Masinagudi	14.16N 75.34 E	11	[c]	11	[e]	4	[f]	5	#	dom	P
India N.	Delhi	28.37N 77.13 E	16	[c]	9	[e]	7	[f]	2	#	<i>dom</i>	P
Pakistan	Islamabad region	33.54N 72.15E	28	[c]	18	[e]	7	[f]	3	#	<i>dom</i>	A(1x) P(1x)
Indonesia	Bogor	06.36S 106.47E	11	[a]	1	[e]	1	#	1	#	musc	A
Russia	Moscow	55.45N 37.37E	7	[d]	7	[d]	3	[d]	2	#	<i>musc</i>	A
Romania	Braila Island	45.17N 27.58E	8	[a]	8	#	1	#	1	#	musc	A
Israel	Bet Nir	31.39N 34.52E	14	[a]								
	Keshet, Ortal	33.03N 35.03E			3	[g]			3	#	dom	P
France	Dordogne	44.27N 00.52E	24	[b]								
	Toulouse	43.37N 01.25E			2	#			2	#	dom	P

Table 2. Allozymic frequencies at 17 polymorphic allozymic loci.

LOCUS	Aat 1		Adh		Amy 1		Es 1		Es-2		Es 10		Hbb		Gpd 1		Idh 1						
	80	100	40	100	80	100	94	100	95	98	100	60	100	100	110	120	90	95	100	100	120	125	
N	4	0	1	0.50	0.50	0	1	0	1	0.25	0.38	0.4	1	0	0.75	0.25	0	0	0.88	0.13	0	0	1
Mashhad	10	0.15	0.85	0.45	0.55	0	1	0	1	0.05	0.95	0	1	0	0.95	0.05	0	0.30	0.65	0.05	0	0	1
Birdjand	4	0.63	0.37	0	1	0.50	0	1	0.25	0.75	0	0	1	0.63	0.13	0.25	0.38	0.63	0	0.13	0	0	0.88
Tehran	8	0.37	0.63	0	1	0.19	0.81	0.94	0.06	0.06	0.06	0.9	0.25	0.75	0.63	0.37	0	0	0.75	0.25	0.82	0.12	0.06
LOCUS	Mor 1		Mod 1		Mod 2		Mpi 1		Np 1		Pgm 1		Pgm 2		Sod 1								
	90	100	100	110	120	100	110	120	100	120	80	90	100	60	80	100	80	100	110	100	80	100	100
Mashhad	0	1	0	0.75	0.25	1	0	0	0.50	0.50	0.63	0.25	0.12	0.25	0.75	0	0.25	0.75	0	1	0	0	0
Kakhk	0	0	0	1	0	1	0	0	0	1	0.30	0.45	0.25	0.50	0.50	0	0.10	0.90	0	1	0	0	0
Birdjand	0.5	0.5	0	0.63	0.37	0.63	0.13	0.24	0	1	0.13	0.12	0.75	0.38	0.37	0.25	0	0.75	0.25	0.38	0.63	0.63	0.63
Tehran	0	1	0.37	0.63	0	0.06	0	0.94	0.13	0.87	0.50	0	0.50	0	0.44	0.56	0	1	0	0	0	0	1

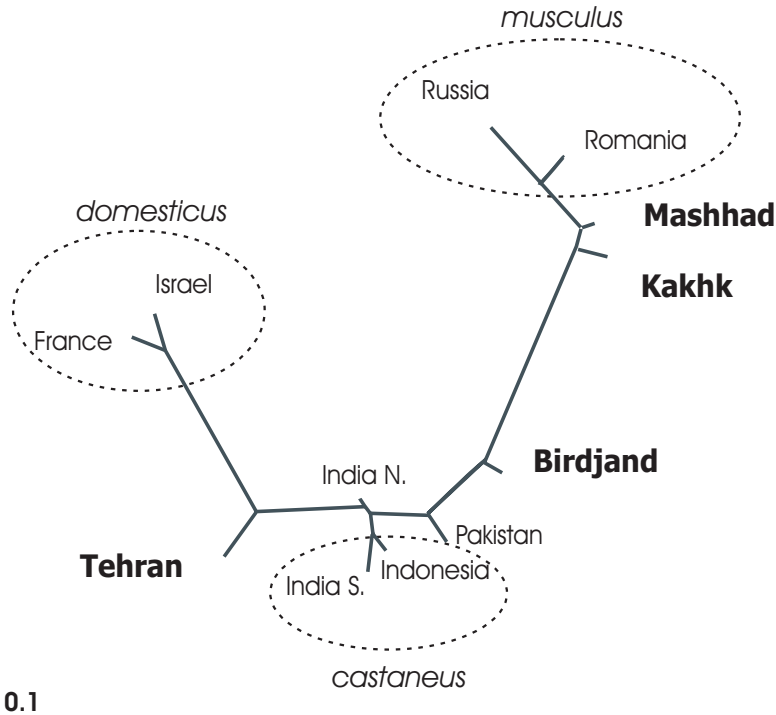


Fig. 2. Neighbour-Joining tree based on Nei's genetic distance at 17 allozymic loci.

suggested by the allozymic data. The same picture was obtained from mtDNA. With the exception of a single oriental haplotype segregating in Shirvan, *musculus* matriline extend as far south as Kakhk. It changes to the oriental matriline further south, in Birdjand and Zabol. Whatever the genetic background, a *musculus*-like Y chromosome as defined by *Zfy-2* is prevalent throughout the eastern Iranian Plateau, although in India and Pakistan the same variant as *M. m. domesticus* is found (Boissinot & Boursot 1997).

Discussion

The allozymic make-up of the Kakhk sample is intermediate between almost pure *musculus* in the north and the central populations in the south, although it is closer to the former than the latter. This raises the question as to whether or not the present day situation corresponds to a primary intergradation formed from an ancestral population in the supposed species cradle that progressively colonised NE Iran. In this scenario of successive bottlenecks, a series of local populations stemming from a southern centre of dispersal would be derived from each other and then subsequently diverge in the very places where they are now. This would imply the existence of a limited level of gene flow, in order to establish and maintain a more or less smooth clinal genetic transition between the archetypal populations in the SE and the *musculus* populations found in the region of Mashhad and further north. Alternatively, it could correspond to a zone of secondary contact similar to the hybrid zones that exist between *domesticus* and *musculus* in Europe (Raufaste et al. 2005 and references therein) and

the Caucasus (Mezhzherin et al. 1998, Orth et al. 1996), or between *castaneus* and *musculus* in Japan (Yonekawa et al. 1986) and China (Moriwaki et al. 1990).

With the limited evidence we have at present, it may be difficult to distinguish between the two hypotheses because their predicted genetic correlates are not very different unless there is a clear signature of a secondary contact in the form of a narrow hybrid zone. Although this is not the focus of the present report, a similar situation might exist further west between *domesticus*- and central populations as already noted by Din et al. (1996) and exemplified by the position of the Tehran sample (Fig. 2). However, a single sample taken from a big city is not definitive proof of a naturally occurring hybrid zone, since mice are known to be prone to man-mediated secondary admixture and long range dispersal. In NE Iran where we have a somewhat larger geographical coverage, we incline for this last hypothesis, since most of the transition occurs between Birdjand and Kakhk which are only 150 km apart. Considering the size of the region this is quite abrupt and is compatible with the existence of a secondary hybrid zone. If, on the contrary, the transition is due to the existence of a progressive primary genetic differentiation, one has to explain its persistence in the face of gene flow. The time required for the divergence between the oriental and *musculus* matriline (several 100 000 years, (She et al. 1990)) is quite sufficient to homogenise genotypes at the scale of the region covered since no major geographical barriers between Birdjand and Mashhad exist or have existed in the past. Further sampling should be performed to further document the nature of this genetic transition

If a secondary contact scenario in the NE of Iran proves to be correct, the exact route taken by the ancestors of the *musculus* population cannot be deduced from the limited data set presented here (see below). Pregel et al. (1998) also found *musculus* matriline in the north of Afghanistan (Maimana, Mazar-e-Sharif, Faizabad (Fig. 1)), with oriental ones in the centre and on the eastern border of this country (Herat, Kandahar, Kabul). So, a secondary contact is likely to exist in Afghanistan as well. From a geographical point of view, the large Harirud river valley is the only major connection at an altitude below 1000 m between the Iranian Plateau and the steppes of Central Asia and is the only significant interruption in the continuous mountain cordilleras that ranges from Caucasus in the west to the easternmost end of the Himalayas in China. As such it is bound to be a very important migration avenue for faunal exchanges between the north and south sides of the Asian continent. This may well correspond to the route taken by the first *M. m. musculus* that invaded the whole northern part of the continent. The mice from NE Iranian plateau would therefore be direct descendents of the ancestral *M. m. musculus*, which eventually came into secondary contact with *M. m. domesticus* elsewhere, like further west in Transcaucasia and elsewhere. This view has recently been challenged by Milishnikov et al. (2004), who proposed from a compilation of allozymic surveys a reversed scenario in which the primary differentiation of *M. m. musculus* occurred in Transcaucasia from where it migrated into Central Asia. In this case it would have entered into NE Iran from the north via the Harirud valley. However, this scenario strongly depends on a correct assessment of the ancestral and secondarily introgressed polymorphisms found in Transcaucasia that we were not able to infirm or confirm with the available data. The cradle of the *M. m. musculus* subspecies therefore, still remains an open subject, as well as the more general question of the geographical origin of the common ancestor of all subspecies evoked in the introduction.

From a taxonomical point of view, one should note that in the dendrogram of Fig. 2 the relative position of the samples we referred to as “central” (N India and Pakistan, to which

Birdjand should be now added) is somewhat different from that of the *bona fide castaneus* mice from Indonesia and S India. Moreover, the type specimen for *M. m. bactrianus* is from Kandahar (Afghanistan) (Marshall 1977), this means that a thorough comparison with other samples from this area, including more markers and phenotypic characters is thus needed before one can decide which subspecific denomination(s) the so-called central populations really deserve.

Acknowledgements

The authors wish to thank B. Dod for useful discussion and corrections on the manuscript, this work was supported in part by the CNRS funding to UMR 5171 and Iranian Research Council grant No.294 to J. D.

LITERATURE

- Bayona-Bafaluy M. P., Acin-Perez R., Mullikin J. C., Park J. S., Moreno-Loshuertos R., Hu P. Q., Perez-Martos A., Fernandez-Silva P., Bai Y. D. & Enriquez J. A. 2003: Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Research* 31: 5349–5355.
- Boissinot S. & Boursot P. 1997: Discordant phylogeographic patterns between the Y chromosome and mitochondrial DNA in the house mouse : selection on the Y chromosome? *Genetics* 146: 1019–1034.
- Bonhomme F., Anand R., Darviche D., Din W. & Boursot P. 1994: The House Mouse as a Ring Species? In: Moriwaki K. (ed.), *Genetics in Wild Mice: Its Application to Biomedical Research. Japan Scientific Societies Press, Tokyo: 13–23.*
- Bonhomme F., Catalan J., Britton-Davidian J., Chapman V. M., Moriwaki K., Nevo E. & Thaler L. 1984: Biochemical diversity and evolution in the Genus *Mus*. *Bioch. Genet.* 22: 275–303.
- Boursot P., Auffray J.-C., Britton-Davidian J. & Bonhomme F. 1993: The Evolution of House Mice. *Ann. Rev. Ecol. Syst.* 24: 119–152.
- Boursot P., Din W., Anand R., Darviche D., Dod B., Von Deimling F., Talwar G. & Bonhomme F. 1996: Origin and radiation of the house mouse : mitochondrial DNA phylogeny. *J. Evol. Biol.* 9: 391–415.
- Britton-Davidian J., Nadeau J. H., Croset H. & Thaler L. 1989: Genic differentiation and origin of Robertsonian populations of the house mouse (*Mus musculus domesticus* Ruddy). *Gen. Res.* 53: 29–44.
- Darvish J. 1995: [Identification of sub-species of *Mus musculus* L. (Muridae, Rodentia) of Khorasan province by multivariate analysis]. *Appl. Ent. Phytopath.* 62: 11–12 (in Persian).
- Din W., Anand R., Boursot P., Darviche D., Dod B., Jouvin-Marche E., Orth A., Talwar G. P., Cazenave P.-A. & Bonhomme F. 1996: Origin and radiation of the house mouse: Clues from nuclear genes. *J. Evol. Biol.* 9: 519–539.
- Felsenstein J.: 1993. Phylogeny inference package. Version 3.5c. *Department of Genetics, University of Washington, Seattle.*
- Gündüz I., Tez C., Malikov V., Vaziri A., Polyakov A. V. & Searle J. B. 2000: Mitochondrial DNA and chromosomal studies of wild mice (*Mus*) from Turkey and Iran. *Heredity* 84: 458–467.
- Marshall J. T. 1977: A Synopsis of Asian Species of *Mus* (Rodentia Muridae). *Bulletin of the American Museum of Natural History* 158: 175–220.
- Mezhzherin S. V., Kotenkova E. V. & Mikhailenko A. G. 1998: The House mice, *Mus musculus* s.l., hybrid zone of Transcaucasus. *Z. Säugetierkd.* 63: 154–168.
- Milishnikov A. N., Lavrenchenko L. A. & Lebedev V. S. 2004: Origin of the house mice (superspecies complex *Mus musculus* sensu lato) from the Transcaucasia region: A new look at dispersal routes and evolution. *Russ. J. Genet.* 40: 1011–1026.
- Moriwaki K., Sagai T., Shiroishi T., Bonhomme F., Chenghui W., Xinquao H., Meilei J. & Zhengan W. 1990: Mouse subspecies differentiation and H-2 polymorphism. *Biol. J. Linn. Soc.* 41: 125–139.
- Nei M. 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.
- Ohtsuka H., Oyanagi M., Mafune Y., Miyashita N., Shiroishi T., Moriwaki K., Kominami R. & Saitou N. 1996: The presence / absence polymorphism and evolution of the p53 pseudogene in the genus *Mus*. *Mol. Phyl. Evol.* 5: 548–556.

- Orth A., Auffray J. C. & Bonhomme F. 2002: Two deeply divergent mitochondrial clades in the wild mouse *Mus macedonicus* reveal multiple glacial refuges south of Caucasus. *Heredity* 89: 353–357.
- Orth A., Lyapunova E., Kandaurov A., Boissinot S., Boursot P., Vorontsov N. & Bonhomme F. 1996: L'espèce polytypique *Mus musculus* en Transcaucasie. *C. R. Biologies* 319: 435–441.
- Page R. D. M. 1996: TreeView: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosciences* 12: 357–358.
- Prager E., Boursot P. & Sage R. 1997: New assays for Y chromosome and p53 pseudogene clines among East Holstein house mice. *Mamm Genome* 8: 279–281.
- Prager E. M., Orrego C. & Sage R. D. 1998: Genetic variation and phylogeography of central Asian and other house mice, including a major new mitochondrial lineage in Yemen. *Genetics* 150: 835–861.
- Prager E. M., Tichy H. & Sage R. D. 1996: Mitochondrial DNA sequence variation in the eastern house mouse, *Mus musculus*: comparison with other house mice and report of a 75-bp tandem repeat. *Genetics* 143: 427–446.
- Raufaste N., Orth A., Belkhir K., Senet D., Smadja C., Baird S. J. E., Bonhomme F., Dod B. & Boursot P. 2005: Inferences of selection and migration in the Danish house mouse hybrid zone. *Biol. J. Linn. Soc.* 84: 447–459.
- Sage R. D., Atchley W. R. & Capanna E. 1993: House mice as models in systematic biology. *Syst. Biol.* 42(4): 523–561.
- She J. X., Bonhomme F., Boursot P., Thaler L. & Catzeflis F. 1990: Molecular phylogenies in the genus *Mus*: comparative analysis of electrophoretic scnDNA hybridisation and mtDNA RFLP data. *Biol. J. Linn. Soc.* 41: 83–103.
- Tanooka H., Sasaki H., Shiroishi T. & Moriwaki K. 2001: p53 pseudogene dating: Identification of the origin of laboratory mice. *Gene* 270: 153–159.
- Yonekawa H., Gotoh O., Tagashira Y., Matsushima Y., Shi L., Cho W., Miyashita N. & Moriwaki K. 1986: A hybrid origin of Japanese mice “*Mus musculus molossinus*”. In: Potter M., Nadeau J.H. & Cancro M.P. (eds), *The Wild Mouse in Immunology. Current Topics in Microbiology and Immunology* 127: 62–67.